DOCUMENT RESUME

ED 315 430 TM 014 446

TITLE	A Glossary of Measurement Terms. ERIC Digest.
INSTITUTION	American Institutes for Research, Washington, DC.; ERIC Clearinghouse on Tests, Measurement, and Evaluation, Washington, DC.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO	EDO-TM-89-1
PUB DATE	Aug 89
CONTRACT	RI-88062003
NOTE	3p.
PUB TYPE	Information Analyses - ERIC Information Analysis Products (071) -- Reference Materials - Vocabularies/Classifications/Dictionaries (134)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Definitions; Glossaries; *Measurement; *Testing; *Vocabulary
IDENTIFIERS	ERIC Digests

ABSTRACT

A glossary of terms commonly used in discussing measurement is presented. Terms, arranged in alphabetical order, contain definitions as they pertain to the measurement field. Terms defined are: (1) achievement test; (2) age norms; (3) average; (4) battery; (5) ceiling; (6) criterion-referenced test; (7) diagnostic test; (8) domain-referenced test; (9) grade equivalent; (10) informal test; (11) inventory; (12) item; (13) norm; (14) normal curve equivalent; (15) norm-referenced test; (16) objective percent correct; (17) percentile; (18) percent score; (19) performance test; (20) published test; (21) rating scales; (22) raw score; (23) reliability; (24) screening; (25) specimen set; (26) standardized test; (27) standard scores; (28) stanine; and (29) validity. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ERIC Clearinghouse on Tests, Measurement, and Evaluation Digest EDO-TM-89-1 August 1989

A Glossary of Measurement Terms

ED315430

A

achievement test -- an objective examination that measures educationally relevant skills or knowledge about such subjects as reading, spelling, or mathematics.

age norms -- values representing typical or average performance of people of age groups.

average -- a statistic that indicates the central tendency or most typical score of a group of scores. Most often average refers to the sum of a set of scores divided by the number of scores in the set.

B

battery -- a group of carefully selected tests that are administered to a given population, the results of which are of value individually, in combination, and totally.

C

ceiling -- the upper limit of ability that can be measured by a particular test.

criterion-referenced test -- a measurement of achievement of specific criteria or skills in terms of absolute levels of mastery. The focus is on performance of an individual as measured against a standard or criteria rather than against performance of others who take the same test, as with norm-referenced tests.

D

diagnostic test -- an intensive, in-depth evaluation process with a relatively detailed and narrow coverage of a specific area. The purpose of this test is to determine the specific learning needs of individual students and to be able to meet those needs through regular or remedial classroom instruction.

domain-referenced test -- a test in which performance is measured against a well-defined set of tasks or body of knowledge (domain). Domain-referenced tests are a specific set of criterion-referenced tests and have a similar purpose.

G

grade equivalent -- the estimated grade level that corresponds to a given score.

I

informal test -- a nonstandardized test that is designed to give an approximate index of an individual's level of ability or learning style; often teacher-constructed.

inventory -- a catalog or list for assessing the absence or presence of certain attitudes, interests, behaviors, or other items regarded as relevant to a given purpose.

item -- an individual question or exercise in a test or evaluative instrument.

N

norm -- performance standards that is established by a reference group and that describes average or typical performance. Usually norms are determined by testing a representative group and then calculating the group's test performance.

TM014446

American Institutes for Research, 3333 K Street, NW Suite 300, Washington, DC 20007 (202) 342-5060

normal curve equivalent -- standard scores with a mean of 50 and a standard deviation of approximately 21.

norm-referenced test -- an objective test that is standardized on a group of individuals whose performance is evaluated in relation to the performance of others; contrasted with criterion-referenced test.

O

objective percent correct -- the percent of the items measuring a single objective that a student answers correctly.

P

percentile -- the percent of people in the norming sample whose scores were below a given score.

percent score -- the percent of items that are answered correctly.

performance test -- designed to evaluate general intelligence or aptitudes. Consists primarily of motor items or perceptual items because verbal abilities play a minimal role.

published test -- a test that is publicly available because it has been copyrighted and published commercially.

R

rating scales -- subjective assessments made on predetermined criteria in the form of a scale. Rating scales include numerical scales or descriptive scales. Forced choice rating scales require that the rater determine whether an individual demonstrates more of one trait than another.

raw score -- the number of items that are answered correctly.

reliability -- the extent of which a test is dependable, stable, and consistent when administered to the same individuals on different occasions. Technically, this is a statistical term that defines the extent of which errors of measurement are absent from a measurement instrument.

S

screening -- a fast, efficient measurement for a large population to identify individuals who may deviate in a specified area, such as the incidence of maladjustment or readiness for academic work.

specimen set -- a sample set of testing materials that are available from a commercial test publisher. May include a complete individual test without multiple copies or a copy of the basic test and administration procedures.

standardized test -- a form of measurement that has been normed against a specific population. Standardization is obtained by administering the test to a given population and then calculating means, standard deviations, standardized scores, and percentiles. Equivalent scores are then produced for comparisons of an individual score to the norm group's performance.

standard scores -- a score that is expressed as a deviation from a population mean.

stanine -- one of the steps in a nine-point scale of standard scores.

V

validity -- the extent to which a test measures what it was intended to measure. Validity indicates the degree of accuracy of either predictions or inferences based upon a test score.

This publication was prepared with funding from the Office of Educational Research and Improvement (OERI), U.S. Department of Education, under contract R-88-062003. The opinions expressed in this report do not necessarily reflect the position or policy of OERI or the Department of Education. **Permission is granted to copy and distribute this ERIC/TM digest.**